USDA Service Center Initiative
Geospatial Data Acquisition, Integration and Delivery
Business Re-engineering Project

Data Themes - Outline - Soils (SSURGO & NASIS)
Prime Farmland & HEL

I.   Acquisition

A. Data Source
1. Producer Information
a.   Name

The soils information in this theme is a combination of spatial data derived from the Soil Survey Geographic Database (SSURGO) and an extraction of tabular data from the National Soil Information System (NASIS).

The prime farmland shapefile is derived from selected elements from the Mapunit table in the MS Access NASIS database. These elements are joined to the soil spatial data using the mapunit symbol as the joinitem. The HEL classifications is added as an attribute field.

National Cartography and Geospatial Center produce SSURGO
The most detailed level of soil mapping done by NRCS. SSURGO digitizing duplicates the original soil survey maps. This level of mapping is designed for use by landowners, townships, and county natural resource planning and management.

Natural Resources Conservation Service produces NASIS
Designed to manage and maintain soil data from collection to dissemination for the National Cooperative Soil Survey.

b.   Location of Headquarters

Natural Resources Conservation Service
P.O. Box 2890
Washington, DC 20013

National Cartography and Geospatial Center
USDA - Natural Resources Conservation Service
P.O. Box 6567
Fort Worth, TX 76115

Information Technology Center
Natural Resources Conservation Service
USDA
2525 Redwing Road
Fort Collins, CO 80526

National Soil Survey Center
USDA - Natural Resources Conservation Service
Federal Building, Room 152
100 Centennial Mall, North
Lincoln, NE 68508-3866

Telephone: (402) 437-4132 FAX: (402) 437-5336

    c.   Internet Address

http://www.ftw.nrcs.usda.gov/soils_data.html
http://www.ftw.nrcs.usda.gov/ssur_data.html

http://nasis.nrcs.usda.gov/

http://www.statlab.iastate.edu/soils/nssc/

2. Publisher Information
    a.   Name

SSURGO data is published by USDA-NRCS-NCGC. It is available through the National Geospatial Data Clearinghouse. Technical information is available from the National Soil Survey Center

NASIS data is available through USDA-NRCS-ITC.

    b.   Location of Headquarters

National Cartography and Geospatial Center
USDA - Natural Resources Conservation Service
P.O. Box 6567
Fort Worth, TX 76115
(800) 672-5559 Telephone (Ordering Information)
(402) 437-5423 Telephone (Tech Support)
(817) 509-5469 FAX

Information Technology Center
Natural Resources Conservation Service
USDA
2525 Redwing Road
Fort Collins, CO 80526

National Soil Survey Center
USDA - Natural Resources Conservation Service
Federal Building, Room 152
100 Centennial Mall, North
Lincoln, NE 68508-3866
Telephone: (402) 437-4132 FAX: (402) 437-5336

    c.   Internet Address

http://www.ftw.nrcs.usda.gov/

http://nasis.nrcs.usda.gov/

http://www.statlab.iastate.edu/soils/nssc/

3. Acquisition Information
    a.   Delivery Media

The map extent for a Soil Survey Geographic (SSURGO) data set is a soil survey area, which may consist of a county, multiple counties, or parts of multiple counties. A SSURGO data set consists of map data, attribute data, and metadata. SSURGO data are available for selected counties and areas throughout the United States and its territories. You will need to know the State Soil Survey Area ID (stssaid) to recognize which files to download.

Digital SSURGO data sets are available via anonymous file transfer protocol (ftp) or CD-ROM (ISO9660).

Currently, NASIS attribute information is extracted for SSURGO map data and is being delivered to the business process reengineering (BPR) pilot sites instead of the SSURGO attribute data. NASIS soils data is being provided to the pilot sites in a format that is different from any format in which soil data has previously been delivered. This new format, referred to as "NASIS Distribution Format 1.0" is based on the structure of a NASIS database. It represents the most current thinking about how soil data should be modeled. This new format will eventually replace both the current SSURGO and FOCS Soils formats.

b.   Download URL

http://www.ftw.nrcs.usda.gov/ssur_ftp.html

c.   Projected Data Availability Schedule

http://www.ftw.nrcs.usda.gov/jpg/ssurgo.jpg (availability status)
ftp://ftp.ftw.nrcs.usda.gov/pub/ssurgo/online98/data/stssaid.txt (textual listing)
http://www.ftw.nrcs.usda.gov/jpg/ssurgo_arc.jpg (archived SSURGO)

B. *Standards Information*
1. Geospatial Data Standard
a.   Standard Name and Steward Information

Draft National Standard for Soil Geographic Data
Soil Data Subcommittee
Federal Geographic Data Committee

b.   Standard Version

November 1996

c.   Standard URL

http://www.fgdc.gov/pub/standards/Soils/soil_std.wp6

http://www.ftw.nrcs.usda.gov/ssurgo.txt (useful README on SSURGO)

2. Metadata Standard
a.   Standard Name and Steward Information

Metadata are compliant to the Federal Geographic Data Committee, 1994, Content Standards for Digital Geospatial Metadata.

b. Description of Metadata Captured
A metadata text file is distributed with each SSURGO data set. The metadata text file provides information on the content, quality, condition, and other characteristics of the data. The sections of metadata include the following:

>   Data Set Identification Information
>   Data Quality Information
>   Spatial Reference Information
>   Status Information
>   Lineage (processing steps)
>   Entity and Attribute Information
>   Distribution Information
>   Data Use Information
>   Metadata Reference Information

c. Metadata Accuracy and Completeness Assessment

The metadata is typically complete. Additional information and a metadata template are available in the geospatial data standard.

C. Acquired Data Structure
1. Geospatial Data Format

a. Format (raster, vector, etc.)

The map data is vector. A SSURGO data set consists of map data, attribute data, and metadata. SSURGO data is collected and archived in 7.5-minute quadrangle units, and distributed as complete coverage for a soil survey area. Soil boundaries ending at quad neatlines are joined by computer to adjoining maps to achieve an exact match.

b. Format Name

SSURGO map data are available in modified Digital Line Graph (DLG-3) optional, Arc INFO coverage and Arc interchange file formats. Attribute data are distributed in ASCII format with DLG-3 map files and in Arc interchange format with uncompressed Arc interchange map files. Metadata are in ASCII format.

c. Data Extent

SSURGO data are available for selected counties and areas throughout the United States and its territories. The map extent for a Soil Survey Geographic (SSURGO) data set is a soil survey area, which may consist of a county, multiple counties, or parts of multiple counties.

d. Horizontal and Vertical Resolution

The accuracy of these digital data is based upon their compilation to base maps that meet National Map Accuracy Standards. The difference in positional accuracy between the soil boundaries and special soil features locations in the field and their digitized map locations are unknown. The location accuracy of soil delineation on the ground varies with the transition between map units.

For example, on long gently sloping landscapes the transition occurs gradually over many feet. Where landscapes change abruptly from steep to level, the transition will

be very narrow. Soil delineation boundaries and special soil features generally were digitized within 0.01 inch of their locations on the digitizing source. The digital map elements are edge matched between data sets. The data along each quadrangle edge are matched against the data for the adjacent quadrangle. Edge locations generally do not deviate from centerline to centerline by more than 0.01 inch.

e.   Absolute Horizontal and Vertical Accuracy

Spatial data meet NRCS standards and specifications for digitizing outlined in Section 647.0508 Digitizing specifications in part 647 Soil Geographic Data Development of the National Soil Survey Handbook.

f.   Nominal Scale

Maps are made at scales ranging from 1:12,000 to 1:63,360. Typically scales are 1:15,840, 1:20,000, or 1:24,000. Scales for new surveys are 1:12,000 or 1:24,000.

g.   Horizontal and Vertical Datum

North American Datum of 1983 and 1927 for DLG files of individual quads
North American Datum of 1983 for Arc coverage and interchange

h.   Projection

Universal Transverse Mercator for DLG, Arc coverage and Arc interchange of individual quads.
Geographic for full Soil Survey Area Arc coverage or Arc interchange

i.   Coordinate Units

Meters for DLG.

Data coordinates for Arc interchange files vary with the geographic extent. Coordinates for quadrangles are in UTM meters. Coordinates for survey area wide Arc coverage are in geographic decimal degrees. Both are referenced to the North American Datum of 1983.

j.   Average Data Set Size

For one Soil Survey Area:

ARC interchanges files average 10 megabytes (compressed)
DLG averages 20 megabytes with 500 kilobytes of attributes (compressed)
The metadata averages 30 kilobytes

k.   Symbology

None

2. Attribute Data Format

a.   Format Name

ASCII files with DLG and Arc interchange format with Arc interchange map files.

b.   Database Size

500 kilobytes of attributes (compressed) per data set

3. Data Model

a. Geospatial Data Structure

DLG Optional, Arc coverage or Arc interchange

b. Attribute Data Structure

ASCII tables or Arc interchange for SSURGO
MS Access tables for NASIS

c. Database Table Definition

Information from:
"Download MS Access 97 NASIS Distribution Format 1.0 Template Database"
http://pestilence.itc.nrcs.usda.gov/pub//NASIS/metadata/ndf10soil.mdb

NASIS Queries and Reports:
Available Interpretations
Erosion Factors
Farmland and HEL Classification
Flooding
Hydric Soil Rating
Land Capability Classification
Prime Farmland
Soil Interpretations
Soil Map Unit Legend
Soil Map Units and Components
Soil Taxonomic Classification
Soil Texture
Static Metadata - Table Columns
Static Metadata - Unique Indexes

NASIS Reports:
Dynamic Metadata - Available Interpretations
Dynamic Metadata - Legend
Dynamic Metadata - Map Unit
Dynamic Metadata - Summary
Dynamic Metadata - Summary - Legend Subreport
Static Metadata - Attributes
Static Metadata - Domains
Static Metadata - Relationships
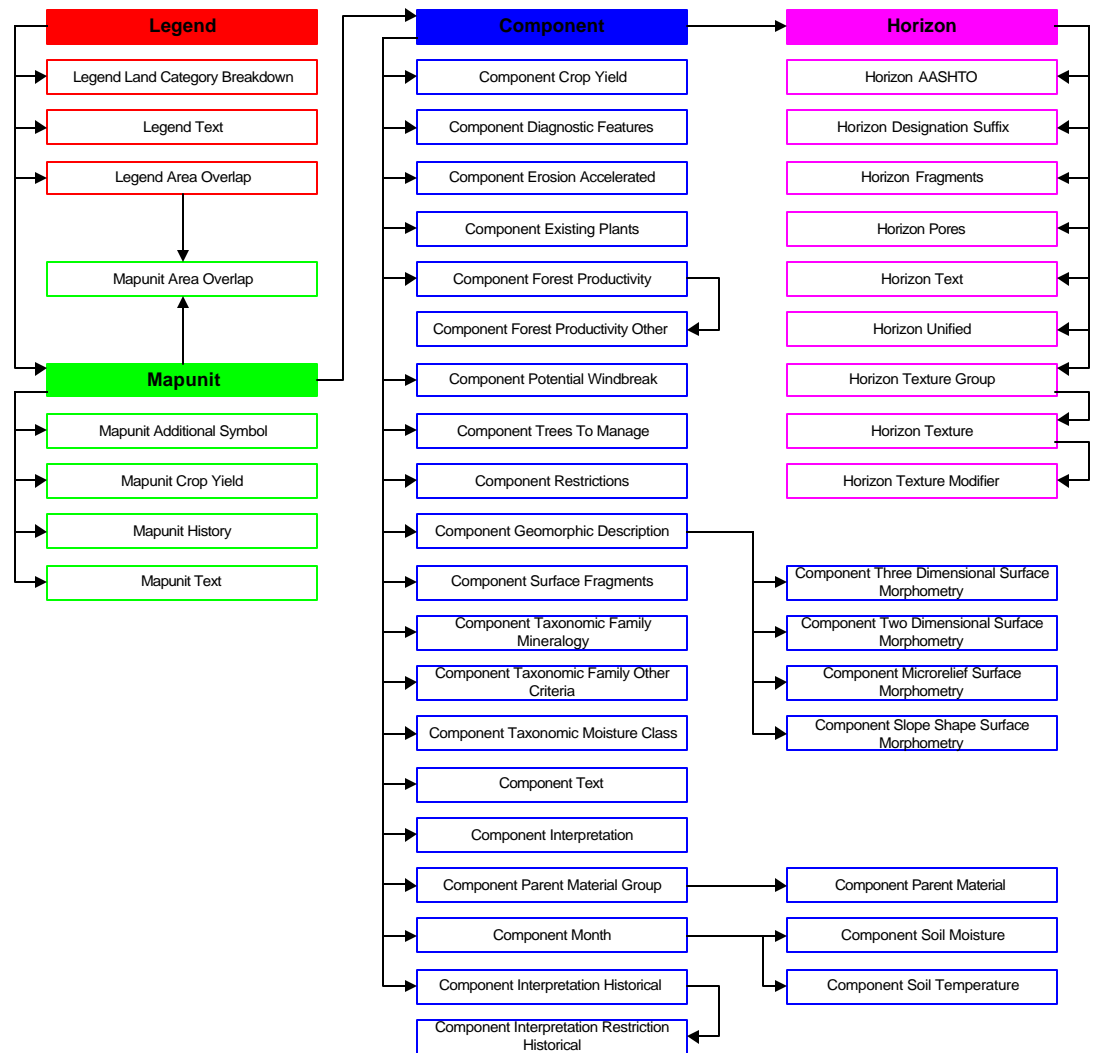Static Metadata - Table Columns
Static Metadata - Unique Indexes

d. Data Relationship Definition

http://pestilence.itc.nrcs.usda.gov/nasis/metadata/df1_0/ndf10traditional.pdf

NASIS Distribution Format 1.0
Data Structure

Legend

- Legend Land Category Breakdown
- Legend Text
- Legend Area Overlap
- Mapunit Area Overlap

Mapunit

- Mapunit Additional Symbol
- Mapunit Crop Yield
- Mapunit History
- Mapunit Text

Component

- Component Crop Yield
- Component Diagnostic Features
- Component Erosion Accelerated
- Component Existing Plants
- Component Forest Productivity
- Component Forest Productivity Other
- Component Potential Windbreak
- Component Trees To Manage
- Component Restrictions
- Component Geomorphic Description
- Component Surface Fragments
- Component Taxonomic Family Mineralogy
- Component Taxonomic Family Other Criteria
- Component Taxonomic Moisture Class
- Component Text
- Component Interpretation
- Component Parent Material Group
- Component Month
- Component Interpretation Historical
- Component Interpretation Restriction Historical

Horizon

- Horizon AASHTO
- Horizon Designation Suffix
- Horizon Fragments
- Horizon Pores
- Horizon Text
- Horizon Unified
- Horizon Texture Group
- Horizon Texture
- Horizon Texture Modifier

- Component Three Dimensional Surface Morphometry
- Component Two Dimensional Surface Morphometry
- Component Microrelief Surface Morphometry
- Component Slope Shape Surface Morphometry

- Component Parent Material
- Component Soil Moisture
- Component Soil Temperature

Shaded Entity - A Fundamental Entity

Clear Entity - A Child of a Fundamental Entity

e. Data Dictionary

The following fields are currently delivered in the polygon .dbf file to the BPR sites. The musym field in the polygon file can be used to link to the NASIS MS Access tables shown in the data relationship definition diagram above:

The following fields are currently delivered in the .dbf files to the BPR sites.

The attribute table of the polygon soils.shp shapefile (.dbf) has the following fields:
- Shape-Spatial data form
- Area-polygon area in square meters
- Perimeter-polygon perimeter in meters
- Name_ -Arc internal feature number
- Name_id-Arc internal feature identifier (Not the map unit ID)
- Musym-alpha or numeric symbol link to MUIR or NASIS tables.

The attribute table Prime farmland shapefile (.dbf) has the same fields as soils.shp plus the following:
- mapunit_sy -unique identifier of the soil map unit in the survey. mapunit_sy is added to provide a check to verify an accurate join with the spatial data.
- mapunit_na -correlated name for the mapunit. The element name was truncated to 10 characters in the process of downloading from Access.
- farmland_c -farmland classification as prime farmland, farmland of statewide importance, farmland of local importance.
- mapunit_he-overall highly Erodible Lands (HEL) class for the mapunit based on the rating of its components for wind and water HEL classification.
- mapunit_ii-record id used with database id (area type, legend, or data mapunit) to uniquely identify the record

The Attribute table of the point soils_pnt.shp, point special features shapefile(.dbf) have the following fields:
- Shape-Spatial data form
- Area-populated by zeros
- Perimeter-populated by zeros
- Name-Arc internal feature number
- Name_id-Arc internal feature identifier
- Featsym-alpha symbol of the point feature
- Symbol-Text name of point feature (not included in all datasets)

The Attribute table of the line soils_lin.shp, line special features shapefile (.dbf) when present have the following fields:
- Shape-Spatial data form
- Fnode-internal number of node where the arc starts
- Tnode-internal number of node where the arc ends
- Lpoly-internal number of the polygon on the left side of the arc
- Rpoly-internal number of the polygon on the right side of the arc
- Length –length of each arc in meters
- Name -internal feature number
- Name_id – internal feature identifier
- Featsym – alpha symbol of the line feature

D. Policies
1. Restrictions
   a. Use Constraints

   The U.S. Department of Agriculture, Natural Resources Conservation Service, should be acknowledged as the data source in products derived from these data.

   This data set is not designed for use as a primary regulatory tool in permitting or citing decisions, but may be used as a reference source. This is public information and may be interpreted by organizations, agencies, units of government, or others based on needs; however, they are responsible for the appropriate application. Federal, State, or local regulatory bodies are not to reassign to the Natural Resources Conservation Service any authority for the decisions that they make. The Natural Resources Conservation Service will not perform any evaluations of these maps for purposes related solely to State or local regulatory programs.

   Photographic or digital enlargement of these maps to scales greater than at which they were originally mapped can cause misinterpretation of the data. If enlarged, maps do not show the small areas of contrasting soils that could have been shown at a larger scale. The depicted soil boundaries, interpretations, and analysis derived from them do not eliminate the need for onsite sampling, testing, and detailed study of specific sites for intensive uses. Thus, these data and their interpretations are intended for planning purposes only. Digital data files are periodically updated. Files are dated, and users are responsible for obtaining the latest version of the data.

   b. Access Constraints

   None

   c. Certification Issues

   For the attribute data derived from NASIS to be used for program activities the state soil scientist must certify it. Also, SSURGO data has been certified that it meets required specifications.

2. Maintenance
   a. Temporal Information

   Digital data files are periodically updated. Files are dated, and users are responsible for obtaining the latest version of the data.

   b. Average Update Cycle

   Periodic

E. Acquisition Cost
1. Cooperative Agreement
   a. Description of Agreement

   None
   The USDA-Natural Resources Conservation Service (former Soil Conservation Service) has been identified by the Federal Geographic Data Committee to coordinate the development of standards for the transfer of soil data. A representative of this agency serves as chairman of the FGDC Soil Data Subcommittee.

b. Status of Agreement

None

2. Cost to Acquire Data

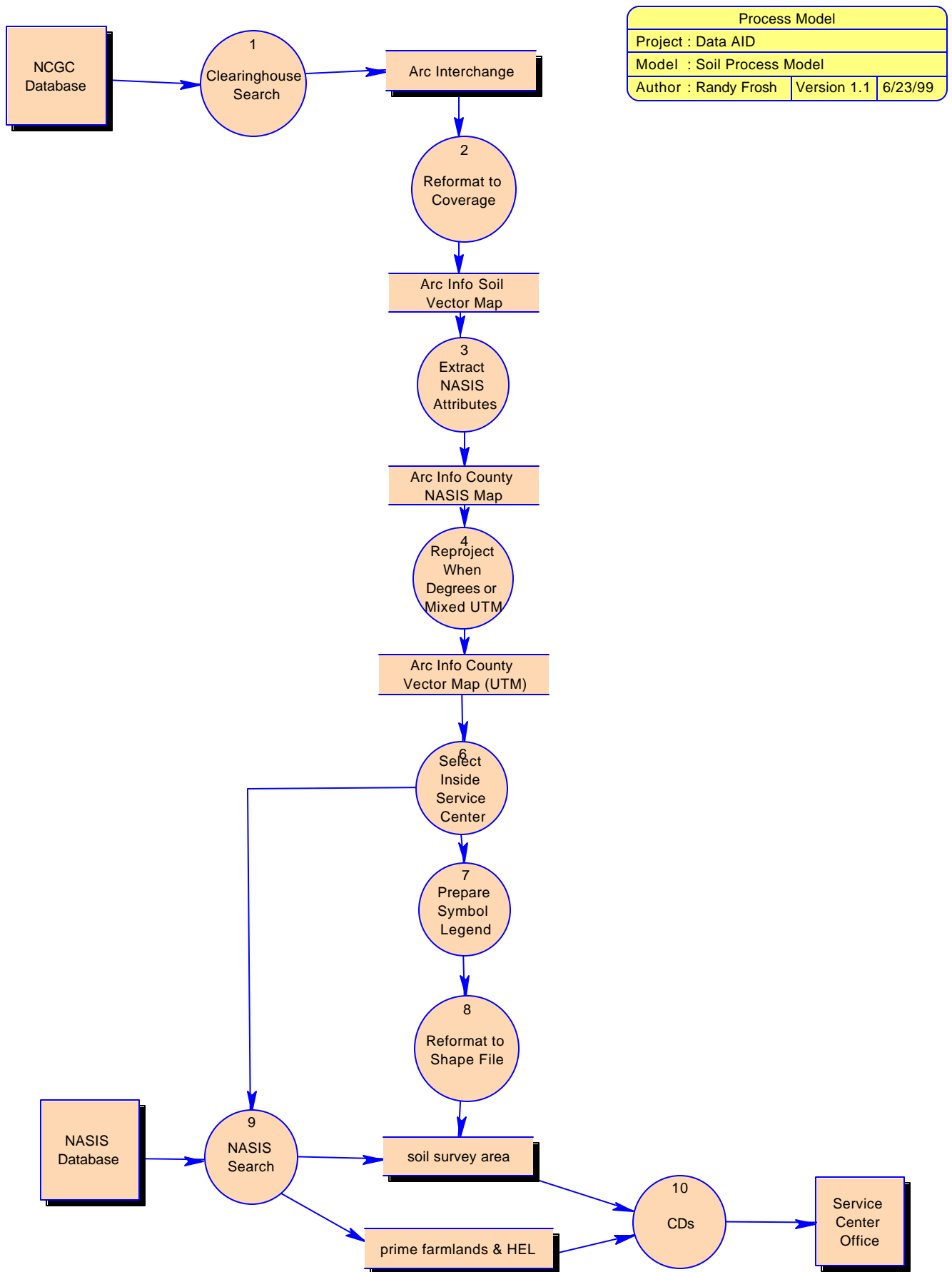FTP is free. $US 50 for CD-ROM of SSURGO map and attribute data.

II. Integration

A. Value Added Process

1. Benefit to the Service Center

Soils data has been identified as a critical layer for several BPR's.

2. Process Model

a. Flow Diagram

Process Model

Process Model		
Project : Data AID		
Model : Soil Process Model		
Author : Randy Frosh	Version 1.1	6/23/99

NCGC Database

1 Clearinghouse Search

Arc Interchange

2 Reformat to Coverage

Arc Info Soil Vector Map

3 Extract NASIS Attributes

Arc Info County NASIS Map

4 Reproject When Degrees or Mixed UTM

Arc Info County Vector Map (UTM)

6 Select Inside Service Center

7 Prepare Symbol Legend

8 Reformat to Shape File

9 NASIS Search

NASIS Database

soil survey area

prime farmlands & HEL

10 CDs

Service Center Office

b. Process Description

Assuming Arc interchange files are used:

- The SSURGO data is downloaded from the NCGC ftp site for the appropriate service center.
- The data is converted to an Arc/Info vector map and attribute tables.
- The attribute fields required to access NASIS data are extracted from the SSURGO data.
- The vector data is re-projected to the dominant UTM zone of the service center when the soil surveys are from more than one UTM zone or when the data was in geographic coordinates by survey wide coverage.
- The data that is inside or crossing the service center boundary is saved.
- The stssaid fields are sent to NASIS so that the MS Access tables can be extracted. Currently, the Soil Survey Area Status (out of date, published, update, in progress) and correlation date is required along with the stssaid field values to extract the NASIS attribute data.
- Using the soils map and its musym field, extract the Prime Farmland fields and HEL classification from NASIS Access database to produce a Prime Farmland map.
- Prepare symbol legend for point and line special features. (The correct symbology is described in exhibit 627-5 NRCS-SOI-37A of the National Soils Handbook   http://www.statlab.iastate.edu/soils/nssh .)
- The saved Arc/Info vector map is converted to a shape and .dbf file(s).
- The NASIS data is downloaded for MS Access.

Please note that currently (Jan-99) the soils survey areas are not combined into one soils map and edge matched for the entire service center for any pilot sites. The data is stored by Soil Survey Area. Nothing in NASIS prevents SSA's from being combined.

3. Technical Issues
   a. Tiling

A SSURGO data set is a soil survey area, which may consist of a county, multiple counties, or parts of multiple counties. This could be changed to a county tiling scheme. However, service center personnel are accustomed to one or more soil survey areas for the service center and perhaps prefer to work with multiple survey maps instead of one soil map for the service center. In addition, the system may be more responsive working with multiple soil survey area maps.

   b. Compression

None

   c. Scale

The soils data is primarily from 1:12,000 or 1:24,000 maps and air-photos. As a result, the data should not be used at a scale larger than this.

   d. Tonal Matching

This is not applicable to vector data.

e. Edge-matching

A SSURGO data set is a soil survey area, wh ich may consist of a county, multiple counties, or parts of multiple counties. For a particular service center, the maps will have to be edge matched if there are multiple soil survey areas that cover the service center area.

4. Quality Control
a. Procedures

This is to be decided. However, it should include checking the NASIS musym against the SSURGO musym and identifying discrepancies. Also, check the match with the imagery and other themes.

b. Acceptance Criteria

This is to be decided.

5. Data Steward
a. Name and Organization

Currently, the data steward for the integrated data is:
National Cartography and Geospatial Center
Natural Resources Conservation Service
US Department of Agriculture
501 Felix Street, Building 23
P. O. Box 6567
Fort Worth, Texas 76115-0567 USA

If the integration procedure can be automated, the steward would optimally remain the same.

b. Responsibilities

Store the SSURGO data and keep it accessible via ftp.

B. Integrated Data Structure
1. Geospatial Data Format
a. Format (raster, vector, etc.)

Vector

b. Format Name

ESRI Shape file

c. Data Extent

Soil survey area.

d. Horizontal and Vertical Resolution

Same as source data.

e. Absolute Horizontal and Vertical Accuracy

Same as source data.

f. Nominal Scale

Same as source data.

g. Horizontal and Vertical Datum

The horizontal datum is the North American Datum (NAD) 83. The vertical datum is mean sea level.

h. Projection

Universal Transverse Mercator (UTM), North American Datum (NAD) 83.

i. Coordinate Units

Meters

j. Symbology
Symbols for point and line special features are selected from available Arc View palettes. There is a need for an NRCS point and line feature palette to be developed in agreement with the Conventional and Special Symbols Legend NRCS-SOILS - 37A.

2. Attribute Data Format
a. Format Name

Dbase V, as part of an ESRI Shape file. MS Access files for NASIS attributes.

b. Database Size

Depends on extent of soil survey area size and variability of soils.

3. Data Model
a. Geospatial Data Structure

Poly Files	
map shp	shp file
map dbf	dbf file
map shx	shx file
map sbn	sbn file
map sbx	sbx file

Point Files	
map shp	shp file
map dbf	dbf file
map shx	shx file
map sbn	sbn file
map sbx	sbx file

Line Files	
map shp	shp file
map dbf	dbf file
map shx	shx file
map sbn	sbn file
map sbx	sbx file

b. Attribute Data Structure

Dbase V, as part of an ESRI Shape file. MS Access for NASIS attributes.

c. Database Table Definition

Standard .dbf file that goes with shape file. For polygon data, the .dbf file includes the additional field musym.

For point data, the .dbf file includes the additional fields featsym and symbol (in some soil survey areas).

For line special feature data, the .dbf file includes the additional field featsym

For Prime Farmland polygons the soils polygon has the additional attribute fields mapunit_sy, mapunit_na, farmland_c, mapunit_he, mapunit_II

    d.   Data Relationship Definition

.dbf to MS Access is planned to be completed as part of Customer Service Toolkit.

    e.   Data Dictionary

Same as in acquired above.

C. Resource Requirements
1. Hardware and Software

To acquire and integrate one set of Soil Sample Area data, a minimum of one UNIX or NT machine with approximately 1-gigabyte of disk is required.

2. Staffing

This is unknown at this time. If the access and integration can be completely automated, it would only require personnel to periodically check the results.

D. Integration Cost
1. Hardware and Software

To reformat, reproject, and subset the dataset a minimum the following is required:
Arc/Info on UNIX or NT platform
ArcView on NT platform
5-gigabyte disk

2. Staffing

This is unknown at this time.

III.   Delivery

A. Specifications
1. Directory Structure
    a.   Folder Theme Data is Stored In

(Version 7)
\Service Center Themes
   \Soils

2. File Naming Convention
    a.   List of Theme Files and The File Naming Convention

\Soils.dbf	\Soil_lin.dbf	\Soil_pnt.dbf
\Soils.shx	\Soil_lin.shx	\Soil_pnt.shx
\Soils.shp	\Soil_lin.shp	\Soil_pnt.shp
\SSA.dbf\	\mlra.dbf	\primefarm.dbf
\SSA.shx	\mlra.shx	\primefarm.shx
\SSA.shp	\mlra.shx	\primefarm.shp

where SSA=soil survey area number. There can be more than one per service center.

Typically there would not be line data as part of SSURGO unless the map went through a digital map finishing process in which case information such as intermittent streams would be included.

Typically, there would not be point data except that nowadays point information is being included for label information and NASIS has geo-reference information for point data such as rock-outcrops.

B. User Information
1. Accuracy Assessment
a. Alignment with Other Theme Geospatial Data

The data is captured at scales varying from 1:12,000 to 1:63,360. There should be some alignment with the ortho-photo layer but this will not be perfect due to the nature of the soils interpretations and the fact that the data is captured at different scales.

b. Content

The data is soils interpretations but can be verified and is the best know soils database.

2. Appropriate Uses of the Geospatial Data
a. Display Scale

The original data source scale or smaller, usually 1:24,000.

b. Plot Scale

The original data source scale or smaller, usually 1:24,000.

c. Area Calculations

As accurate as the source data and capture scale and the algorithm used by ArcInfo/ArcView.

d. Decision Making

Unknown what information is requested here.

C. Maintenance and Updating
1. Recommendations and Guidelines
a. Frequency of Updates

Update the Service Center whenever a Soil Survey Area is updated that is in the Service Center.

b. Location for the Theme Data to be Maintained

Ideally, the data would be extracted from NCGS and the NASIS sites, processed, then stored at the Service Center using the data.

c. Maintenance and Updating Procedures Overview
This depends on the update cycle and whether or not the access and integration can be fully automated.